

Intrasensory Synesthesia in Musical Composition

Fernando Egido, Composer, busevin@gmail.com

Abstract. I will present the concept of Intrasensory Synesthesia and how it has been used in several musical compositions. Using the synesthesia as a model, intrasensory synesthesia tries to make possible the perception of a musical parameter or sound features through another one. In the same way that through synesthesia a stimulus in a sensory can be perceived in a different sensory path. We will introduce several new compositional techniques based on the idea of intrasensory synesthesia. They use cognitive and semiotic knowledge to allow a sound object in a musical flow defined by a sound parameter be perceived in a different one. So instead of hearing colors, we try to build a musical flow that will let us perceive pitches troughs rhythm or the timbre through intensity.

I am going to enumerate the cognitive and semiotic knowledge in which this idea is based. 1) The way in which sound parameter influences the perception of the other ones. 2) The threshold of perceptions of certain musical features. 3) The semiotic strategies based on the music cognition features. 4) The possibility of listening conditioning to focus the musical flow into certain sound features.

Finally, I am going to introduce the concepts of *parametric interdefinition*, and *parametric modulation* as the two main new composition techniques that allow the intrasensory synesthesia to happen. I show how these techniques have been used in three works, *Three Chants for Computer*¹, *Homo Homini Lupus* and *Parametric modulation studies*.

Keywords: intrasensory synesthesia, cognitive musicology, parametric music, semiotics.

1. INTRODUCTION

Can we interpret timbre as pitch in a sound flow in the same way we can taste colors through synesthesia? My answer is yes. From 1960, composers have granted a special interest in the perception of sound as a source of inspiration and knowledge in such a manner that we can consider the perception itself as musical material. Risset (RISSET J. C. 1969) showed us how we can use the peculiarities of perception to create paradoxes in the perception of sound. In his work *Mutations*, he creates the illusion of a never-ending glissando following the Shepard model (SHEPARD. R. 1964) Risset was the creator of a new type of music in which what is perceived is different from what actually is happening using cognitive knowledge.

1.1 Towards the concept of intrasensory synesthesia.

Intrasensory synesthesia is the phenomenon that happens when a parameter or a property of a sense is perceived as another parameter or property in the same sense. In extrasensory synesthesia, we have different senses, each of them associated to a property of objects that is usually perceived via its corresponding sense, but for a neurological reason,

this relation changes and one property of the object is perceived as another different sense. Intrasensory synesthesia encompasses only one sensory system characterized by different properties (timbre, pitch, intensity, and duration); in such a way that into the same sensory system a property of the sound is recognized as if it were another one.

Intrasensory synesthesia is based on two facts. First, the extrasensory synesthesia between auditory systems and sensory system it always happens through one parameter of the sound flow, only.

The Synesthesia evokes a character: a change in frequency induces kinetic sensations that tend to be of amplitude in the case of big intervallic changes, or proximity, in the case of chromatic intervals. (BRAGANÇA, 2008)

Secondly, the neuroscience has established that each sound feature or parameter is processed in a different brain region, so any sound feature can be considered as an independent perceptual system that can interact between them as if they were individual ones.

These data suggest that distinct areas subserve musical meter, tempo, pattern, and duration. In addition, the results confirm that the neural subsystems underlying pitch discrimination are distinct from those subserving

¹ These Works can be listened at <https://busevine.wordpress.com/music/>

the elements of rhythm discrimination. The differences in activated brain areas for non-musicians and musicians likely reflect differences in strategy, perceptual skill, and cognitive representation of the components of musical rhythm a pitch (PARSONS, 2003)

1.2 *Introducing principal concepts.*

The main idea is that the perception of the sound flow or the musical form is relative and it is not directly dependent on the objective properties of the material. The music based in Intracensory Synesthesia it is differentiated from the other kind of music where the musical flow it is not dependent on the properties of materials, as for example pitch in tonal music.

Cognitive psychology deals in part with the problem that what we perceive is not exactly the same as the physical reality of the stimulus. (SHEPARD, 1999)

The sound flow it is no longer dependent on a property of sound working as central. The musical tension, created from the material properties, is substituted by the cognitive tension derived from the processing of the formal information encoded in the parametric relationships. The concept of cognitive dissonance is key to create new formal structures, the tension of the formal structure it is not dependent on the properties of materials that cause dissonance but in the cognition tasks involved in decoding the patterns of the work.

A cognitive dissonance occurs when the brain has established a specific criterion of elements order and an element appears unexpectedly, not the expected one according to the cognitive principle deduced by the cognitive systems. Cognitive dissonances replace traditional dissonances in my music. The flow of tension relaxation occurs by the greater or less quantity of objects that are cognitive dissonances (EGIDO, 2011).

Cognitive dissonance is based on cognitive models instead of the properties of materials to create tension in the musical flow of the sound stream.

There are two groups of techniques used to create music based on Intracensory Synesthesia. The first one uses cognitive knowledge, for example, the form in which one property of sound interferes or influences the perception of a different one. The sound perception is far from being perfect and the form in which properties of the sound are perceived depends on the range of the values of the rest of properties or parameters. In fact, there is a range of all sound properties where all the parameters do not interfere: in the middle range. Classical music is based on this range and all the classical and tonal music use the range in which the parameters do not interfere and perception works well. So we will just do the opposite. We are creating new musical concepts and musical or sound materials that use the

way in which certain properties influence the perception of others using the ranges of these properties so they will influence each other. Afterward, these micromaterials are organized in parametric structures where the musical tension is generated creating an expectation that won't be accomplished, and conditioning listening to interpret the meaning of certain parameters to another via cognitive models.

The second group of techniques used in the Intracensory Synesthesia Music derives from semiotic concepts that encode the meaning of what is happening in one parameter to a different one. This second group of techniques is not only formal and it is also related to cognitive models that allow this to be possible.

1.3 *Using interferences between parameters perception.*

The first composer to use some limit in the threshold of perception of the sound properties was Grisey (GRISEY, 1982). This method of Intracensory Synesthesia consists in using the interferences where certain ranges in the parameters influence how we perceive a different one to create musical concepts. One of the most evident examples of this is the concept of parametric interdefinition that uses the threshold of perception of one parameter to create the illusion that this parameter is perceived by means of a different one. *This means that we perceive a change of a sound property of a sound object by changing a different parameter.*

The parametric interdefinitions are based on the idea that a parameter can be interdefined with another. An example of this in everyday life is the interdefinition of the distances through the time it takes us.

Basically, there are two different kinds of parametric interdefinitions. One is semiotical (the meaning of a parameter is relative to another.) and the other one is based on the thresholds of perception (liminal).

The Semiotical one is based on the idea that the meaning of a parameter, i.e. the implication of information or meaning loads is carried out by the difference between two values of a parameter being relative to another parameter. The semiotical interdefinitions are carried out by the significance space which is a semantic function that maps a particular difference of values in a property (parameter), of a musical object to a meaning based on another parameter. This type of parametric interdefinition is widely used in my first three studies.

The procedure to create music composition driven by Intracensory Synesthesia is: first, we select some cognitive knowledge (for example, some particular perception interference between parameters

perception). Second, we convert this into a musical material that permits the listener to have the sensation that a feature of sound is semiotically determined by another one (for example, parametric interdefinition, parametric modulation or parametric morphing). These materials are called parametric materials and are a set of microelements that form the basic materials of the architecture used by building the work. It must be underlined that these materials are not real things in the same way that a melody is not a real thing. A melody as a real thing is only a succession of frequencies and durations. The same happens for these materials. This paper only focusses on the micromaterials. For a further detailed explanation of large organization of these materials, please consult: (EGIDO, 2011).

2. ENUMERATION OF COGNITIVE KNOWLEDGE USED CREATE PARAMETRIC INTERFEFINITION TECHNIQUES ²

2.1 Intensity regarding Pitch

The parametric interdefinition between pitch and intensity can be done using the Fletcher diagram. (FLETCHER, MUNSON, 1933)³

Pitch determines that for the same sound pressure there is a different intensity perception. For example, the same objective sound pressure can be considered as *ppp* in the bass register but as *mp* in the middle register, so frequency can determine the perception of intensity.

We can use it also to interdefine pitch with timbre and timbre with intensity, given that the harmonics between 500 and 5000 Hz will be highlighted.

2.2 Pitch regarding Intensity

The pitch perception and the interval perception are influenced by intensity; higher sound is perceived higher at high intensity. So the pitch sensation depends on intensity (STEVEN, 1935).

2.3 Pitch regarding Octavation

The octavation perception is influenced by the pitch. In the high register, we need more than the double to perceive the upper octave, for example, in order to perceive the following octave from a sound of 2000 Hz we need a sound of 4600 Hz instead of 4000.

This phenomenon is in more detail explained by a diagram that relates mels with pitch. (STEVEN, DAVIS, 1945)

These three parameter influences have been used extensively in *Three Chants for Computer* especially in the chant number one and in *Homo homini lupus*. The use of the inharmonic timbres in combination with octavation in high register and in contrast with the octavation in the middle register creates a great confusion about pitch, timbre and fundamental recognition that leads to an ambiguous and entangled perception that creates the sensation of musical motion when in reality nothing is happening from a classical perspective (as there is no harmonic movement).

2.4 Pitch regarding timbre

The fundamental partial recognition depends on timbre. If the partials have the same distance between them the fundamental recognition is easier. If the timbre spectrum is inharmonic the stronger partial will be recognized as fundamental. Sound filtering can be used to alter fundamental recognition.

These results suggest that pitch discrimination is facilitated by the presence of the harmonic partials. (TERVANIEVI, 2003).

The modulatory and carrier frequencies have different roles in timbre perception regarding AM Modulation.

Carrier frequency determines the rate of stimulation along the basilar membrane. While the modulation rate determines the periodicity-mediated pitch. (DOWLING, HARWOOD, 1986)

So, using AM modulation we can get the recognition of a very different timbre changing only the carrier frequency. This was intensively used in *Homo homini lupus* to produce an ambiguity in the timbre recognition between voice timbres and synthesized sounds. In the same way, if the modulatory frequency is multiple of the carrier the voice recognition is easier, so we can use these two parameter influences in such a way that the timbre perception (and voice recognition) is determined by a process carried in the frequency domain.

Two sounds with the same modulation frequency in each ear but different carrier frequency are perceived merged. If the two sounds have the same carrier but different modulation frequency they are perceived as different timbres to each ear. This was used also in *Homo homini lupus* to provide and stereo image based on a slightly different perception associated to each ear created and ambiguous space image in which the difficulty of the recognition was determined by the differences in the presentation of

² A very interesting review of all the forms in which same parameters influence the perception of the others can be found in the chapters IX- X – XI of (LEIPP, E. 1980). This list does not exhaust all the possible interferences between parameter perception, only the important ones from a compositional point of view are quoted.

³ A short description of the phenomena can be found at) (DOWLING, HARWOOD, 1986).

the carrier and modulation frequency in a different form in each ear.

2.5 Pitch regarding formants

For the vowel recognition is more effective the formants absolute peak frequency than the relative one. However, if the fundamental moves in the same proportion that formants the vowel recognition is easier. (SLAWSON, 1968).

This was used intensively in *Homo homini lupus* in which, in certain parts of the work, the ambiguity, between voice and synthesized sounds, was provided by changing the proportions between the formants, in such a way that with just relatively tiny changes the difficulty in vowel recognition increased a lot.

2.6 Time regarding timbre

The brain needs a certain amount of time for recognizing timbre. Below this threshold no timbre recognition is possible. This threshold is around 20 ms.

At the same time, formant analysis for speech requires high time resolution whereas tonal patterns require good frequency resolution.

Temporal and spectral resolution are inversely related, so that improving temporal resolution can only come at the expense of degrading spectral resolution and vice versa (ZATORRE, 2003)

Time may interfere with almost any other sound feature by means of the forms. Firstly, as we have seen before, if the sound is short enough to be under the threshold of time needed to be recognized we can make that the object is perceived in one form or another controlling its duration. Regarding moving objects (for example a glissando) we can also control how this object is recognized by controlling the velocity of change.

These two techniques were used especially in *Homo homini lupus*. The synthesis parameters change in different velocities in every section of the work so, the recognition of the sound object depends on how a velocity of parametric changes makes much easier or difficult the recognition, taking into account the oppositions.

2.7 Missing fundamental

The brain can perceive a fundamental which actually does not exist. If all the upper harmonic partials are separated by the same constant frequency, this is called residual phenomena. For this phenomenon to happen at least three upper partials are required and they must be between 50 – 200 Hz.

2.8 Timbre with rhythm

A pulse of short sound grains will be perceived as a train of pulses or as pitched sound with a concrete timbre depending on the velocity of the grains. The threshold that permits this is between 16-20 Hz, depending on each individual. From here on pulses will be perceived as a pitched timbre sound. This has been used especially in the second Chant of *Three Chants for Computer* as we will see later.

3. UTILIZATION OF PARAMETRIC INTERDEFINITION AND PARAMETRIC MORPHING

As I mentioned earlier, the first chant of *Three Chants for Computers* was built using several parametric interdefinitions. This set of techniques interdefine timbre with several aspects of pitch. And pitch with other features of pitch itself. It explores several forms in which the harmonic spectra interfere with pitch perception and vice versa, and the way in which octavation affects pitch perception and timbre. A special technique of time interdefinition with timbre was also used, in it, a very short fragment of a bowed metal sheet sound is stretched by Phase Vocoder techniques. Thanks to that the irregular and inharmonic partial, it appears as a melody or as a timbre controlling the stretching of time; so, depending on the scale of time the same objective sound object is perceived with a very different feature perspective.

In the first chant are only used parametric interdefinitions using parameter perception interferences.

The most important material derived from the parametric interdefinition is parametric morphing in which time perceived event is converted into a timbre perceived object using a certain threshold of perception and the influence of the pitch perception in the timbre. This was used in the second chant.

Parametric morphing is a special case of parametric interdefinition in which three parameters are implicated instead of two.

For example, a really fast pulse sequence will begin to be perceived as a timbre from a certain threshold of velocity (Pitches), being possible to parameterize a given object in such a way that varying a single parameter it is perceived as defined by a different parametric centrality. There is a threshold of a parameter from which a particular object is perceived in one way or another. Thus, two different sound object properties can be interdefined through a third one, (for example: timbre, time, frequency) so that altering only one parameter (frequency), below certain threshold of frequency (pulse velocity), the sound object is perceived from the centrality of a property (time or rhythm) and above this threshold it is perceived from the centrality of the third one

(timbre). We can perform parametric modulations by altering a single parameter.

An example can be taken from the second chant of *Three Chants for Computer*. Using a FOF granular synthesis (CLARKE, RODET, 2003). We start from a tiny grain, representing a pulse (time) that is repeated at a very low frequency (0.5 Hz) then, gradually, as we increase the pitch of the grain, it seems that the pulse accelerates until it becomes very fast. From 16

Hertz on, it starts crossing the threshold from which it is perceived as an accelerating pulse and it begins being perceived as an upward glissando with a definite voice timbre. The sound object passes from being recognized as a temporary object to a timbral object by altering only its frequency.

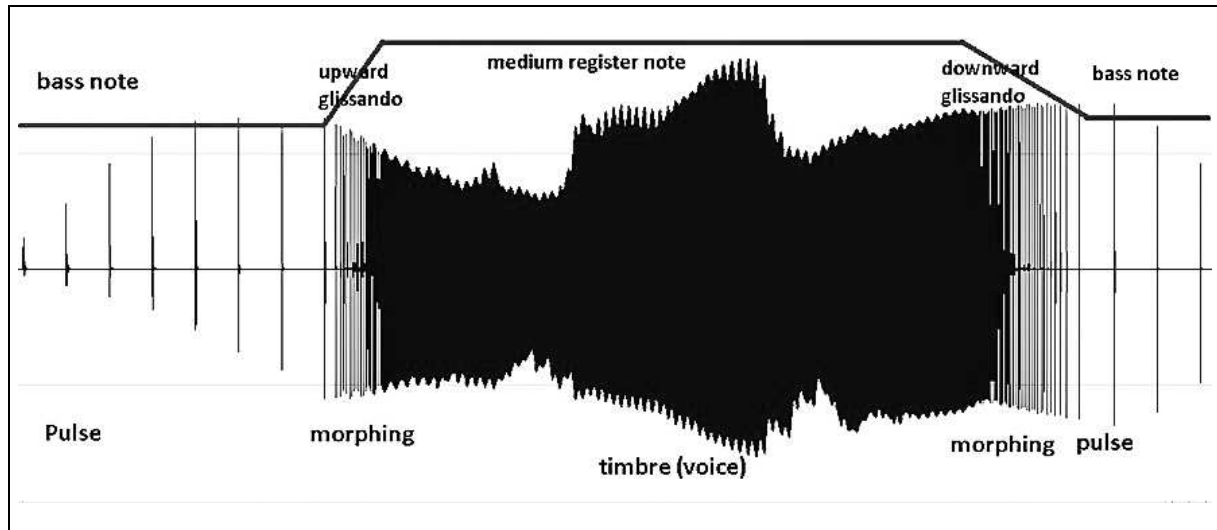


Figure 1, an example of parametric morphing, where a musical object is perceived as timbral or rhythmical depending on pitch.

4. SEMIOTIC STRATEGIES TO PROVIDE INTRASENSORY SYNESTHESIA

The basic idea is that we can condition listening to associate the meaning of certain sound parameter to another one in the same manner in which we can associate visual stimulus as colors to auditory stimulus as frequencies. Therefore, a set of values in the domain of a sound property can be correlated to another set of values of another sound feature, in such a way that the first set work as a sign of the second one. This arbitrary correlation can be only formal, but it works better if we use some kind of cognitive or psychological knowledge to make this correlation to be more real, so listening is conditioned to correlate the sign with its meaning.

For example, we can create a pattern in which the change of all the sound parameters is correlated to only one (determinant parameter) in such a way that the change of the parameters is associated to the determinant parameter. The first section of *Parametric neutralization studies* is divided into three subsections. In all the three sections intensity is always correlated to the change of one parameter while the other remains neutralized (they do not change), so in the first subsection frequency is associated to intensity in such a way that when we have the maximum intensity we have a very dense cluster that evolves to only one note correlated to

the diminishing of intensity. In the second subsection dynamic is correlated to a different parameter: time timbre, so while noise is associated to maximum intensity, sinusoidal timbre is associated to low intensity, frequencies and rhythm remain neutralized. In the third subsection, the rhythmic complexity is associated to maximum intensity and rhythm simplicity.

This way listening is conditioned to associate each parameter change to intensity. So any change is understood by its correlation with intensity. So we can condition the listening to understand the formal information that happens in several parameters to a determiner parameter (that works as a sign of the other ones).

The Intrasensory Synesthesia driven Music works by creating several formal data streams, each of them codified by a different parameter but correlated in such a way that listening is conditioned to associate the meaning of the overall process each time to a different parameter.

We have seen a basic and simple technique of listening conditioning, but to provide large musical forms we need to create some more abstract semiotical concept that is the signification space that is a semiotical application that correlates the meaning of what is happening in one parameter to another.

For example, in the same way, that in tonal music each note of a tonality has a different role. (I, IV, V grades are important, the rest are neuter) providing formal information according to its role. In the Cognitive-Parametric Music, each parameter has a different role and changes every time that a parametric modulation is performed. When the roles of the parameters change we have a parametric modulation that means, from a intrasensory synesthetic point of view, that the relation between the parameters changes, creating cognitive dissonance in a way that the parameters are now correlated in a different way and the brain has to do an especial cognitive task to find the new patterns that make the new correlation between parameters.

Another way of performing parametric modulations and listening conditioning is using the influence that the numbers of elements in a scale have in the facilities to recognize the process. According to (Miller, G. A. 1956), across various sensory modalities, the number of stimuli along a given dimension people can categorize consistently is typically 7 ± 2 With more than seven people begin to hear two or more pitches as falling in the same category.

So, if we use a scale of different stimuli numbers for each sound parameter the listening will focus on the process of those that are below 9 or seven elements. In this way, if we have a scale of 12 pitches and a scale of 5 timbres the patterns associated to timbre will be easier to recognize. Using this knowledge, we can perform parametric modulations associating a different number to the stimuli scale to each parameter and changing the distribution of this scale along parameters.

Table 1. Parametric modulation using different stimuli scale.

Parameter	Timbre parametric tonality	Intensity parametric tonality
Durations	1 stimulus	1 stimulus
Pitch	1 stimulus	1 stimulus
Intensity	1 stimulus	5 stimulus
Timbre	5 stimuli	1 stimuli

This can be performed using statistical distribution curves. We can perform a parametric modulation by changing the distribution of the probability distribution around parameters.

Table 2. Parametric modulation using different distribution probability functions.

Parameter	Timbre parametric tonality	Intensity parametric tonality
Durations	uniform	uniform
Pitch	uniform	uniform
Intensity	uniform	gaussian
Timbre	gaussian	uniform

These two techniques can be improved by using symmetries versus asymmetries.

If values of the determiner parameter (the one that has been conditioned to work as a sign) are distributed asymmetrically (with different distances between the values) the pater associated to this parameter will be recognized much easier. So we can also use this to improve the parametric modulations.

Table 3. Parametric modulation using symmetries versus asymmetries.

Parameter	Timbre parametric tonality	Intensity parametric tonality
Durations	symmetrical	symmetrical
Pitch	symmetrical	symmetrical
Intensity	symmetrical	asymmetrical
Timbre	asymmetrical	symmetrical

These techniques have been used intensively in my Works *Parametric neutralization studies* and *Parametric modulation studies*.

REFERENCES

- BRAGANÇA, G. F. F. (2008), "A Sinestesia e a Construção de Significação Musical". Bello Horizonte Escola de Música - UFMG
- CLARKE M. AND RODET, X., (2003) " Real-time FOF and FOG synthesis in MSP and its integration with PSOLA". International Computer Music Conference, 2003
- DOWLING AND HARWOOD, (1986) " Music Cognition". Academic Press.1986, Orlando Florida.
- EGIDO F. (2011), "Towards an Aesthetics of Cognitive-Parametric Music", Auto edition.
- FLETCHER, H. AND MUNSON, W. A., (1933) "Loudness, its definition, measurement and calculation". Journal of the Acoustical Society of America, 5, 82.
- GRISEY G. (1982), "LA MUSIQUE : LE DEVENIR DES SONS" in "Écrits ou L'invention de la musique spectral", Édition Établie Par Guy Lelong. Collection Répercussions, éditions MF, Paris.
- LEIPP, E. (1980) " Acoustique et Musique " E. Masson. Third edition. Paris.

- MILLER G. A (1956), *"The magical number seven, plus or minus two, some limits of our capacity for processing information"* Psychological Review, 1956, 63, 81-97.
- PARSONS L. M, (2003) *"Exploring the Functional Neuroanatomy of Musical Performance, perception, and Comprehension"* In PERETZ I., ZATORRE R. *"The Cognitive Neuroscience of Music"*. OXFORD UNIVERSITY PRESS.
- RISSET J. C. (1969), *"Pitch Control and Pitch Paradoxes Demonstrated with Computer-Synthesized Sounds"* Journal of the Acoustical Society of America, 46, 88.
- SHEPARD R. (1999). *"Pitch Perception and Measurement"* In COOK P. R. *"Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics"*. MIT PRESS
- SHEPARD R. (1964). *"Circularity in Judgments of Relative Pitch"* Journal of the Acoustical Society of America, 35, 2346-2353.
- SLAWSON, A. W. (1968). *"Vowel quality and musical timbre as function of spectrum envelope and fundamental frequency"*. Journal of the Acoustical Society of America, 43, 87-101.
- STEVEN, S.S., (1935) *"The relation of pitch to intensity"*. Journal of the Acoustical Society of America, 6, 150-154.
- STEVEN, S.S. AND DAVIS H., (1945) *"Hearing"*. Wiley, New York.
- TERVANIEVI M. (2003) *"Musical sound processing: EGG and MEG evidence"* In PERETZ I., ZATORRE R. *"The Cognitive Neuroscience of Music"*. OXFORD UNIVERSITY PRESS.
- ZATORRE, G.J., (2003) *"Neural Specializations for tonal Processing"* In PERETZ I., ZATORRE R. *"The Cognitive Neuroscience of Music"*. OXFORD UNIVERSITY PRESS.